СОВРЕМЕННЫЕ НАКОПИТЕЛИ НА ЖЕСТКИХ МАГНИТНЫХ ДИСКАХ (часть 1)

Вадим Морозов, Сергей Яценко (г. Ростов-на-Дону)

Приступаем к публикации большой статьи, из которой Вы узнаете все об устройстве, эксплуатации, дефектах накопителей на жестких магнитных дисках и методах восстановления информации на них. Материал может быть интересен как специалистам, обслуживающим компьютерную технику, так и «продвинутым» пользователям, желающим получить представление о наиболее сложной детали своего компьютера.

На сегодняшний день наиболее распространенные операционные системы являются дисковыми, т.е. запускаются с жесткого диска компьютера. Результаты работы любых прикладных программ, будь то бухгалтерия или дизайн, также сохраняются на диске. Поэтому надежность, быстродействие и информационная емкость накопителей на жестких магнитных дисках (НЖМД) являются определяющими параметрами качества любого компьютера.

УСТРОЙСТВО НАКОПИТЕЛЕЙ

Конструктивно НЖМД (или «винчестер») состоит из механической части – герметизированного блока (HDA) и платы электроники (PCB). В гермоблоке размещена вся механика винчестера, а также микросхема предварительного усилителя/коммутатора. На плате электроники установлены микросхемы, управляющие механическими узлами, кодированием/декодированием данных с магнитного носителя и приемом/передачей информации через внешний интерфейс. Плата электроники размещается за пределами гермоблока, – как правило, в нижней его части. В некоторых моделях винчестеров, например известной серии Barracuda фирмы Seagate, контроллер закрыт металлической крышкой, защищающей электронику от наводок и повреждений.

Механическая часть

Основу всей конструкции составляет литой металлический корпус, предохраняющий точную механику от воздействия окружающей среды. Внутренний объем корпуса заполнен очищенным от пыли воздухом. Сборка механики осуществляется в цехе, где на один кубический метр воздуха приходится менее ста пылинок (т.н. чистой зоне класса 100).

В гермоблоке размещается пакет магнитных дисков, насаженный на шпиндель двигателя, магнитные головки с системой позиционирования и предусилитель/коммутатор, обеспечивающий усиление сигнала с головок и переключение между ними.

В корпусе гермоблока имеется отверстие, закрытое плотным фильтром, обеспечивающее выравнивание внутреннего давления воздуха с атмосферным давлением. (Именно через это отверстие внутрь накопителя может попасть вода.) При вращении дисков создается поток воздуха, который циркулирует внутри корпуса и, постоянно проходя через еще один фильтр, дополнительно очищается от пыли.

Магнитный диск представляет собой круглую пластину из алюминия (в редких случаях из керамики или специального стекла), поверхность которой обработана по максимально возможному классу точности, за исключением зоны парковки (если таковая имеется). Из-за экстремальной чистоты обработки поверхностей дисков и магнитных головок, они могут прилипать друг к другу вследствие сил молекулярного притяжения. Для предотвращения этого явления зона контакта головок и диска имеет искусственную шероховатость – специальную лазерную насечку.

Чтобы придать поверхности диска магнитные свойства, ее покрывают ферросплавом на основе оксида хрома или вакуумным напылением наносят слой кобальта. Такое покрытие имеет высокую твердость и существенно большую износостойкость, чем у ранних моделей винчестеров, в которых пластины покрывались слоем не очень стойкого магнитного лака на основе окиси железа.

Вращение дисков осуществляет специальный 3-фазный *электродвигатель*. Статор двигателя содержит три обмотки, включенные звездой с отводом посередине, а ротор – постоянный секционный магнит из редкоземельных металлов. Для обеспечения малого биения на высоких оборотах в двигателе используются специальные подшипники, которые могут быть как шариковыми, так и более совершенными – жидкостными (вместо шариков в них используется специальное масло, поглощающее ударные нагрузки, что увеличивает долговечность двигателя). Жидкостные подшипники имеют более низкий уровень шума и почти не выделяют тепло во время работы. Скорость вращения двигателя в современных накопителях с интерфейсом IDE составляет 5400 или 7200 об./мин, с интерфейсом SCSI – 10 000 или 15 000 об./мин.

Магнитная головка представляет собой сложную конструкцию, состоящую из множества деталей. Эти детали настолько малы, что изготавливаются методом фотолитографии (как микросхемы). Рабочая поверхность керамического корпуса головки отполирована с такой же высокой степенью точности, что и диск. Привод головок представляет собой плоскую катушку, помещенную между полюсами постоянного магнита и закрепленную на конце рычага, вращающегося на подшипнике. На другом конце рычага находится подвес с магнитными головками. Сам подвес подпружинен с определенным усилием, что позволяет головкам «лететь» на высоте в десятые доли микрометра над поверхностью диска.

Привод, перемещающий блок головок, получил название *Voice Coil («звуковая катушка»)* – из–за прямой аналогии с устройством электродинамического громкоговорителя. Обмотку позиционера окружает постоянный магнит. При подаче на катушку тока определенной величины и полярности рычаг поворачивается в соответствующую сторону с определенным ускорением. Изменяя ток в обмотке, можно устанавливать головки в любое положение на траектории в плоскости диска.

Для фиксации головок в нерабочем состоянии (в зоне парковки) используются специальные защелки. Наибольшее распространение получили два типа защелок – магнитные и воздушные. Магнитная защелка представляет собой маленький постоянный магнит, который в положении парковки головок притягивает железную пластинку – флажок, установленный на корпусе обмотки. Воздушная защелка (или воздушный замок) также фиксирует позиционер в парковочной зоне, не давая ему перемещаться. С началом вращения дисков создаваемый ими воздушный поток отклоняет «парус» воздушной защелки и освобождает систему позиционирования.

В гермоблоке находится *предусилитель/коммутатор сигналов* с магнитных головок. Это необходимо, чтобы уменьшить наводки на гибкий шлейф (плоский кабель), идущий от головок к основной плате. К этому же шлейфу подключена «звуковая катушка», а иногда и шпиндельный двигатель. Однако в большинстве накопителей питание на двигатель подается через отдельный шлейф, поэтому подключение гермоблока к плате управления, как правило, осуществляется двумя разъемами.

Плата электроники

Схемотехника современных накопителей отличается применением небольшого числа микросхем разной степени интеграции. Структурная схема электронной части приведена на рис. 1.

Электрическая схема современных накопителей, как правило, реализована на 6 интегральных микросхемах. Однако в последних моделях появились БИС контроллеров системы, совмещающие канал чтения/записи, микроконтроллер на основе RISC–процессора и контроллер ввода/вывода. Таким образом, количество ИМС на плате электроники НЖМД сократилось до четырех:

- контроллер системы, включающий канал чтения/записи, контроллер ввода/вывода и микроконтроллер;
- микросхема ЭППЗУ, содержащая микропрограмму;
- микросхема управления двигателем и приводом головок;
- микросхема ОЗУ.

Дальнейшее увеличение степени интеграции микросхем накопителя затрудняется принципиально разными энергетическими режимами работы его функциональных узлов.

Первым совмещенным контроллером системы, использованным в накопителях, была БИС, разработанная фирмой Cirrus Logic. К сожалению, недостаточно отработанная методика применения этой БИС привела к частым отказам накопителей серий MPF3xxxAT и MPG фирмы Fujitsu.

Микроконтроллер содержит процессорное ядро, реализованное на основе RISC–архитектуры. После подачи питающего напряжения или активизации сигналом внешнего интерфейса, схема сброса подает сигнал RESET на микроконтроллер, который, отрабатывая микропрограмму, записанную в ПЗУ, выполняет самодиагностику, очищает рабочую область памяти данных, программирует дисковый контроллер и другие микросхемы, находящиеся на внутренней шине НЖМД.

После этого микроконтроллер опрашивает устройства накопителя и, если нет сигналов об аварий–

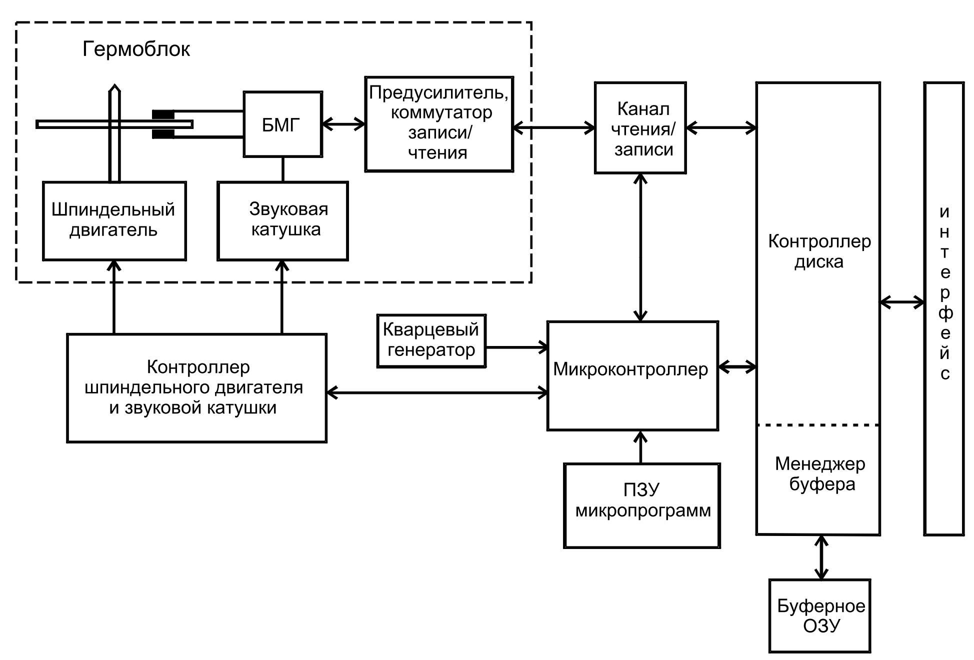


Рис. 1. Структурная схема накопителя

ных ситуациях, выдает сигнал запуска двигателя. Следующий этап работы микропрограммы – проведение внутреннего теста ОЗУ буфера данных и команд, контроллера диска и состояния входных сигналов со стороны внешнего интерфейса. Далее микроконтроллер, измеряя период следования импульсов фазных обмоток, ожидает, пока двигатель не наберет номинальную скорость вращения. После этого он выдает команду на перемещение магнитных головок в зону, где записана служебная информация, и пересылает считанные данные в буферное ОЗУ. Теперь микроконтроллер готов к приему сигналов с внешнего интерфейса. В этом режиме команда, пришедшая от центрального процессора компьютера, порождает цепочку действий, в которой участвуют все элементы НЖМД.

Канал чтения/записи состоит из предусилителя/коммутатора (расположенного в гермоблоке), схемы чтения, схемы записи и генератора опорных временных последовательностей.

Предусилитель имеет несколько каналов, каждый из которых соединен со своей головкой. Переключение осуществляется по сигналам микроконтроллера. Предусилитель содержит переключатель тока записи и сенсор ошибок, который формирует соответствующий сигнал, если в головке произошло короткое замыкание или обрыв.

Интегрированный канал записи/чтения в режиме записи осуществляет прием данных с дискового контроллера синхронно с тактовой частотой записи, производит их кодирование, осуществляет предкоррекцию и передает в усилитель записи. В режиме чтения сигнал с выхода предусилителя/коммутатора поступает на схему АРУ, затем, проходя через программируемый фильтр, схему адаптивной компенсации и импульсный детектор, преобразовывается в последовательность данных, которые поступают на контроллер диска для дальнейшей обработки и передачи через внешний интерфейс.

Контроллер диска – наиболее сложный элемент накопителя – определяет скорость обмена данными между накопителем и центральным процессором компьютера.

Контроллер диска имеет четыре порта, с помощью которых он подключается к ЦП, микроконтроллеру, буферному ОЗУ и каналу ввода/вывода (интерфейсу). Дисковый контроллер представляет собой автомат, управляемый микроконтроллером. Со стороны ЦП доступны только стандартные регистры контроллера. Программирование дискового контроллера производится на этапе инициализации при помощи микроконтроллера, при этом выполняется настройка методов кодирования данных, выбирается полином исправления ошибок чтения, устанавливается режим гибкого или жесткого разбиения на сектора и т.д.

Менеджер буфера данных является функциональной частью дискового контроллера и управляет буферным ОЗУ, объем которого в современных накопителях составляет от 512 Кбайт до 8 Мбайт. Менеджер разбивает все пространство буферного ОЗУ на секторы. Специальные регистры, доступные со стороны микроконтроллера, содержат начальные адреса этих секторов. Когда ЦП осуществляет обмен данными с одним из них, со стороны канала чтения/записи возможен обмен данными с другим сектором буферного ОЗУ. Таким образом распараллеливаются процессы чтения/записи данных и обмена с ЦП.

Контроллер двигателя управляет 3-фазным двигателем, его программирование осуществляется при помощи микроконтроллера. Существуют три режима управления двигателем: запуск, ускорение и стабильное вращение. Рассмотрим режим запуска. При включении питания подается сигнал сброса на микроконтроллер, который, выполняя инициализацию, записывает в регистр контроллера двигателя команду на запуск. Контроллер формирует сигналы переключения фаз, при этом шпиндельный двигатель вращается с небольшой скоростью и генерирует э.д.с. самоиндукции при переключении обмоток. Контроллер двигателя обнаруживает этот сигнал и передает микроконтроллеру, который использует его для измерения частоты вращения шпинделя. В режиме ускорения микроконтроллер увеличивает частоту переключения фазных обмоток, постоянно измеряя скорость вращения двигателя. Когда номинальная скорость достигнута, контроллер переводит двигатель в режим стабильного вращения.

В этом режиме микропроцессор вычисляет время одного оборота двигателя, основанное на сигнале разности фаз, и подстраивает скорость вращения. После вывода магнитных головок из зоны парковки скорость вращения управляется сигналом серворазметки.

Контроллер привода головок формирует управляющий ток перемещения и стабилизации позиционера на заданном треке. Значение тока вычисляется микроконтроллером на основе оцифрованного сигнала рассогласования положения головки относительно трека (Position Error Signal, PES). Значение тока в цифровом виде подается на ЦАП. Аналоговый сигнал с выхода ЦАП усиливается и подается на «звуковую катушку» позиционера головок.

Служебная информация

Служебная информация необходима для функционирования самого НЖМД и, как правило, скрыта от пользователя. Служебную информацию можно разделить на четыре основных типа:

- серво-информацию, или серворазметку;
- формат нижнего уровня;
- резидентные микропрограммы (рабочие программы);
- таблицы конфигурации и настройки.

Серворазметка необходима для работы сервосистемы привода магнитных головок НЖМД. Именно по серворазметке осуществляется их позиционирование и удержание на дорожке. Сервисная разметка записывается на диск в процессе производства через специальные технологические окна в корпусе собранного гермоблока. Затем окна заклеиваются шильдиками, на которых написано «DO NOT OPEN». Запись осуществляется собственными головками накопителя при помощи специального высокоточного прибора – серворайтера. Перемещение позиционера головок осуществляется специальным толкателем серворайтера по калиброванным шагам, которые намного меньше межтрековых интервалов.

Рабочие программы (микрокод) управляющего микроконтроллера представляют собой набор программ, необходимых для работы НЖМД. К ним относятся программы первоначальной диагностики, управления вращением двигателя, позиционирования головок, обмена информацией с дисковым контроллером, буферным ОЗУ и т.д. В большинстве моделей накопителей рабочие программы размещаются во внутреннем ПЗУ («прошивке») микроконтроллера, в других моделях используется внешнее ЭППЗУ. В некоторых моделях НЖМД часть рабочих программ хранится в служебной зоне магнитного диска, а в ПЗУ микроконтроллера записаны программы инициализации, позиционирования и начальный загрузчик для считывания рабочих программ с магнитного диска в ОЗУ. Вследствие того, что программные модули перегружаются в ОЗУ при запуске накопителя, они получили название «резидентных».

Производители жестких дисков размещают часть микропрограмм на магнитном носителе не только для экономии объема ПЗУ, но и для возможной оперативной коррекции кода, если в процессе производства или эксплуатации обнаруживаются ошибки. На Интернет-сайтах почти всех производителей НЖМД можно найти утилиты, выполняющие такую модернизацию. Переписать микропрограмму на диске значительно проще, чем перепаивать «прошитые» микроконтроллеры. (По этой причине фирма Western Digital несколько лет назад отозвала большую партию накопителей.)

Формат нижнего уровня. Начало дорожки определяется индексным импульсом. Каждая дорожка разделена на секторы данных и сервометки. Формат каждого сектора содержит поле идентификации, поле данных, синхрозоны и пробелы. В начале каждого сектора расположена синхрозона, которая служит для фазирования и синхронизации строб-импульса выделения данных. Поле идентификации содержит адресный маркер, физический адрес сектора, байт флагов и байты циклического контрольного кода.

В последнее время получил распространение формат без использования идентификаторов. При таком способе расположения информации на дорожке поля идентификации вообще отсутствуют, что увеличивает полезную емкость дисков. Вместо них используется система сервоуправления, располагающая физические секторы на дорожке. Чтение/запись всех секторов производится за один оборот диска, при этом в ОЗУ хранятся образы считанной и записываемой дорожек. Таким образом, для чтения одного сектора в ОЗУ «помещается» вся дорожка целиком, и чтение последующих секторов выполняется не с диска, а из ОЗУ накопителя. Для записи одного сектора происходит чтение дорожки, ее модификация в ОЗУ и запись целой дорожки на диск.

Таблицы конфигурации и настройки накопителей содержат информацию о логической и физической организации дискового пространства. Они необходимы для самонастройки электронной части диска, которая одинакова для всех моделей семейства. Например, в процессе «проектирования» модели с емкостью 80 Гбайт на четырех поверхностях двух пластин автоматически получается «половинная» модель (40 Гбайт на одной пластине или двух поверхностях) или «четвертинка» (20 Гбайт на одной поверхности). Таким образом, производитель может выпускать фактически одинаковые НЖМД разной емкости и стоимости. Кроме того, в «младших» моделях можно использовать детали, которые не могут быть использованы в «старших» моделях. Например, в моделях НЖМД «половинной» емкости используются диски, имеющие дефекты на одной из поверхностей.

Таблицы дефектов. Современная технология изготовления магнитных дисков не позволяет изготовить диски без дефектов. Неоднородность материала носителя, огрехи шлифовки, посторонние примеси при нанесении магнитного слоя и т.д. порождают области поверхности диска, на которых невозможна безошибочная запись и чтение данных.

В накопителях 15-летней давности, например типа ST506/412, таблица дефектных дорожек была напечатана на наклейке, размещенной на корпусе гермоблока, а сам «винчестер» имел запас по емкости. Например, НЖМД типа ST225 (20 Мбайт) на самом деле имел емкость 21,5 Мбайт, т.е. 1,5 Мбайта отводилось под сбойные сектора и дорожки. В современных НЖМД также имеется избыточность, но она доступна только микроконтроллеру. Одна часть избыточной емкости отводится под рабочие программы, таблицы конфигурации, счетчики S.M.A.R.T., заводскую информацию, таблицы дефектов и т.п. Другая часть резервируется для замены сбойных секторов, возникающих в процессе работы «винчестера».

Заполнение таблиц дефектов производится во время внутризаводских испытаний. Номера сбойных блоков помещаются в таблицу. Такая процедура называется скрытием или подгрузкой дефектов (UPDATE DEFECT). Если при работе НЖМД произойдет обращение к скрытому дефектному сектору, оно будет переадресовано резервному сектору. Поэтому все современные накопители, вышедшие с завода-изготовителя, не имеют ни одного дефектного сектора.

Большинство моделей НЖМД имеют две таблицы дефектов: начальную (Primary, или P-List) и растущую (Grown, или G-List). Начальная таблица заполняется на заводе-изготовителе при выполнении внутризаводского тестирования – SELFSCAN (intelligent burn-in). Вторая таблица предназначена для размещения дефектов, возникающих во время эксплуатации. Для этого в списке пользовательских команд практически всех НЖМД существует команда «assign», по которой происходит переназначение сбойного сектора на резервный. Эту команду используют многие тестовые программы, в том числе рекомендованные фирмой-производителем для «лечения» накопителей.

В накопителях фирмы Western Digital реализована система Data Lifeguard, выполняющая переназначение сбойных секторов автоматически, когда к накопителю нет обращений со стороны ЦП. Для этого накопитель тестирует свои поверхности и переносит данные пользователя в резервный сектор, а дефектный сектор помечается соответствующим образом. Такой механизм скрытия дефектов по сути аналогичен процедуре «assign».

В накопителях Fujitsu, Quantum, Maxtor, IBM реализован механизм автоматического скрытия дефектов во время операции записи. Так, если запись данных производится в дефектный сектор, накопитель самостоятельно переадресует обращение в ре-

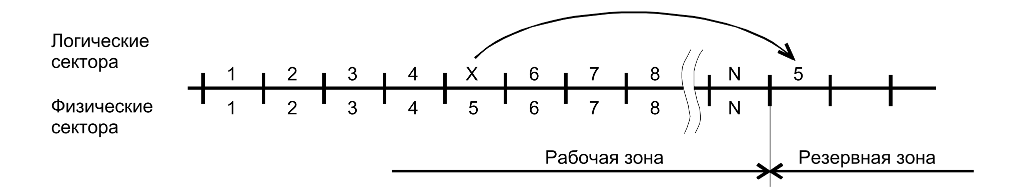


Рис. 2. Метод переназначенного сектора

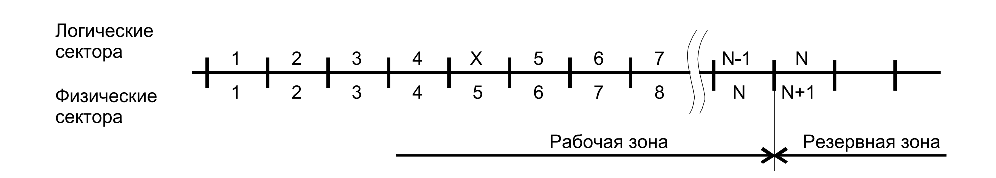


Рис. 3. Метод пропущенного сектора

зервную зону, а дефектный сектор помечается как сбойный и его номер заносится в таблицу (G–List). Из специализированных программ, служащих для скрытия дефектных секторов, можно отметить FUJFMT.EXE для накопителей Fujitsu, WDDIAG.EXE для накопителей Western Digital, ShDiag.exe от Samsung и т.п.

Два механизма скрытия дефектов

При использовании механизма переназначения (assign) накопитель прописывает в поле идентификации дефектного сектора флаг перемещенного сектора и далее в поле данных этого сектора – номер резервного сектора, к которому необходимо обратиться для записи или чтения данных. Как правило, резервным сектором является первый сектор после рабочей области данных пользователя (см. рис. 2).

При обращении к сбойному сектору во время операции чтения/записи данных контроллер читает флаг и адрес переназначения и производит позиционирование в резервную зону, чтобы осуществить чтение/запись данных в хорошем секторе. Дефектные сектора в этом случае «пропадают», но накопитель каждый раз, наткнувшись на сбойный сектор, позиционирует головки в резервную область. Этот процесс сопровождается щелчками и некоторым замедлением работы.

При помощи процедуры «assign» можно скрыть только дефекты, находящиеся в полях данных. Если ошибки связаны с разрушением полей идентификации или поврежденной сервометкой, такие дефекты методом переназначения скрыть нельзя.

Второй механизм скрытия дефектных секторов, реализуемый на заводе-изготовителе, заключается в пропуске дефектного сектора. При использовании такого метода дефектный сектор игнорируется, а следующему за резервным сектором присваивается номер дефектного (и так далее, по цепочке); самый последний сектор сдвигается в резервную зону (см. рис. 3).

Такой метод скрытия нарушает целостность последовательности форматирования нижнего уровня, т.е. система трансляции логической структуры диска LBA (logical block addressing – логическая адресация блоков) в физическую структуру PCHS (physical cylinder, head, sector – физические цилиндр, головка, сектор) должна учитывать дефектные сектора, чтобы избежать обращения к ним. Метод пропуска требует пересчета таблиц транслятора и выполнения команды низкоуровневого форматирования, что делает невозможным его использование без потери данных пользователя. Именно по этой причине данный метод скрытия дефектов выполняется только в специальном (технологическом) режиме работы накопителя при помощи упомянутой выше утилиты FUJFMT.EXE, которая предназначена для скрытия дефектов в накопителях фирмы Fujitsu.

Логическая организация дискового пространства

В современных накопителях значительная часть дискового пространства скрыта от пользователя. Она содержит служебную информацию и резервную область для замены сбойных секторов НЖМД. В нормальном режиме работы скрытая часть доступна только микроконтроллеру. Пользователю отведена рабочая область, называемая логическим дисковым пространством. Именно ее объем указан в характеристиках НЖМД. Обращение к рабочей области, представляющей собой неразрывную цепочку логических секторов, происходит с номера 0 по N (в структуре LBA). Связь между логическим дисковым простран-

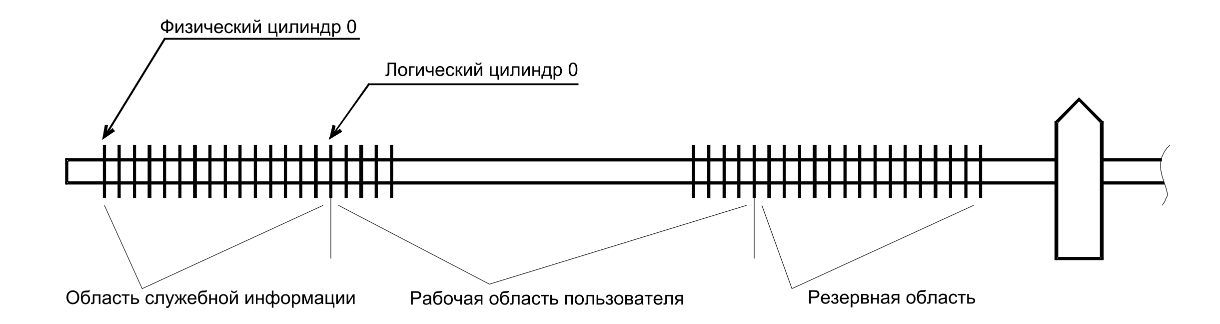


Рис. 4. Логическая организация дискового пространства

ством и физическим форматом осуществляется посредством специальной программы-транслятора, которая учитывает физический формат, зонное распределение и исключенные дефектные сектора и треки.

Доступ к служебной зоне возможен только в технологическом режиме работы накопителя. Для входа в этот режим подается команда-ключ, после чего открывается доступ к дополнительному, технологическому набору команд. При помощи этих команд осуществляются такие операции, как запись/чтение сектора служебной зоны, получение карты расположения модулей и таблиц в служебной зоне, получение таблицы зонного распределения, перевод LBA в PCHS и обратно, запуск формата нижнего уровня, запись/чтение ЭППЗУ и др.

В процессе проектирования НЖМД разработчик определяет необходимую для функционирования накопителя служебную информацию и отводит под нее соответствующее количество цилиндров. Поэтому логическим нулевым цилиндром является первый свободный цилиндр, следующий за последним цилиндром служебной информации (см. рис. 4). Структура дискового пространства может отличаться у различных моделей НЖМД.

Продолжение следует.